



TITLE:

Analysis of the effect of degree correlation on the size of minimum dominating sets in complex networks

AUTHOR(S):

Takemoto, Kazuhiro; Akutsu, Tatsuya

CITATION:

Takemoto, Kazuhiro ...[et al]. Analysis of the effect of degree correlation on the size of minimum dominating sets in complex networks. PLoS ONE 2016, 11(6): e0157868.

ISSUE DATE:

2016-06-21

URL:

<http://hdl.handle.net/2433/216028>

RIGHT:

© 2016 Takemoto, Akutsu. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

RESEARCH ARTICLE

Analysis of the Effect of Degree Correlation on the Size of Minimum Dominating Sets in Complex Networks

Kazuhiro Takemoto¹*, Tatsuya Akutsu²*

1 Department of Bioscience and Bioinformatics, Kyushu Institute of Technology, Iizuka, Fukuoka, Japan,
2 Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Kyoto, Japan

* These authors contributed equally to this work.

* takemoto@bio.kyutech.ac.jp (KT); takutsu@kuicr.kyoto-u.ac.jp (TA)



OPEN ACCESS

Citation: Takemoto K, Akutsu T (2016) Analysis of the Effect of Degree Correlation on the Size of Minimum Dominating Sets in Complex Networks. PLoS ONE 11(6): e0157868. doi:10.1371/journal.pone.0157868

Editor: Naoki Masuda, University of Bristol, UNITED KINGDOM

Received: March 16, 2016

Accepted: June 6, 2016

Published: June 21, 2016

Copyright: © 2016 Takemoto, Akutsu. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information file.

Funding: KT and TA were partially supported by grants from JSPS KAKENHI (numbers 25700030 and 26540125, respectively). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Network controllability is an important topic in wide-ranging research fields. However, the relationship between controllability and network structure is poorly understood, although degree heterogeneity is known to determine the controllability. We focus on the size of a minimum dominating set (MDS), a measure of network controllability, and investigate the effect of degree-degree correlation, which is universally observed in real-world networks, on the size of an MDS. We show that disassortativity or negative degree-degree correlation reduces the size of an MDS using analytical treatments and numerical simulation, whereas positive correlations hardly affect the size of an MDS. This result suggests that disassortativity enhances network controllability. Furthermore, apart from the controllability issue, the developed techniques provide new ways of analyzing complex networks with degree-degree correlations.

Introduction

Controlling complex systems is a central challenge in a wide range of research fields (e.g., molecular networks and brains in biology, the Internet and WiFi communication in telecommunications engineering, and economic science). Since complex systems are often represented as networks, a network perspective is helpful for understanding the control and design of the systems [1, 2], especially on networks with a scale-free property (i.e., a power-law distribution of degree k : $P(k) \propto k^{-\gamma}$) [1].

Liu et al. [3] proposed network controllability characterized by the minimum set of driver nodes (so-called *structural controllability*), which is sufficient to control the whole-system dynamics and is identified based on the maximum matching of the network. They found that network controllability is mainly determined by degree distributions; in particular, they showed that real-world networks, which are generally sparse and heterogeneous, are difficult to control.

It is also important to consider assortativity or degree-degree correlation [4], a structural property widely observed in real-world networks, when discussing the relationship between network structure and system dynamics because assortativity influences the system dynamics. For example, a negative degree-degree correlation (i.e., disassortativity) protects networks against the spread of viruses [5]. Disassortativity enhances synchronization [6], and it promotes observability [7]. Degree-degree correlation indicates the relationship between the node degrees, and it is often defined as the Pearson correlation coefficient of degrees between a connected node pair [4] (i.e., assortative coefficient r_a). In particular, a negative and a positive r_a indicate the disassortativity and assortativity of networks, respectively. Pósfai et al. [8] investigated the effect of degree-degree correlation on network controllability in the context of driver nodes or maximum matching, and they showed that positive and negative degree-degree correlations tend to increase the number of driver nodes (i.e., they reduce network controllability); however, degree correlation coefficients have quadratic or no relationship with the number of driver nodes, depending on the nature of the underlying correlations.

However, more careful examinations are required to conclude the effect of degree-degree correlation because network controllability is also investigated in the context of minimum dominating sets (MDSs) [9–16] whose approach is conceptually similar to the controlling link dynamics [17]. In this context, network controllability is defined as the proportion of the MDS to network size. In particular, heterogeneous networks are known to show relatively small MDS sizes (i.e., they are easy to control) [10, 11]; however, this result does not conflict with the conclusion obtained by Liu et al. [3] because the MDS-based approach employs a stronger assumption that each node in an MDS can control its outgoing edges independently, although it also considers structural controllability under linear dynamics. Apart from the traditional control theory based on continuous/discrete dynamics, the MDS is important in the control of discrete engineering systems. The MDS is a well-known concept in graph theory and has already been applied to the design and/or control of various types of discrete engineering systems, which include mobile ad hoc networks (MANET) [18–20], transportation routing [21], and computer communication networks [21, 22].

Molnár et al. studied the relation between degree-degree correlation and MDS size [13]. They employed a probabilistic approach developed by Alon and Spencer and analyzed the expected size of a random dominating set (RDS) and a cutoff dominating set (CDS) under degree-degree correlations, where RDS is a dominating set (DS) obtained by randomly selecting nodes with some specified probability and adding the nodes not dominated by these nodes, and CDS is a dominating set obtained by selecting nodes above a degree threshold and adding the nodes not dominated by these nodes (see Ref [13] for details). Note that RDS and CDS provide the upper bounds of the size of an MDS (i.e., they are dominating sets but are not necessarily the minimum ones), with CDS giving better estimates in general than RDS does. Although their analytical methods provide accurate estimates of the size of an RDS and a CDS, they also showed experimentally that there is a large gap between the sizes of an MDS and an RDS (respectively, a CDS). Therefore, an accurate estimate of the size of an MDS under degree-degree correlation remains an unsolved problem.

In this study, therefore, we investigated the impact of degree-degree correlation on the size of an MDS both analytically and numerically, and showed that disassortativity, which indicates that high-degree nodes tend to connect to low-degree nodes, reduces the size of an MDS. To fill the known gap between the sizes of an MDS and an RDS, we developed two novel techniques: (i) decomposition of assortative (respectively, disassortative) networks, (ii) use of recursive probabilistic estimation methods for analyzing the size of an MDS in k -regular random graphs and (h, k) -regular random bipartite graphs. The former one is based on the proper understanding of assortative and disassortative network structures. The latter one is considered

as a recursive application of the analysis technique used in Ref [13]. Although we applied both of them to the analysis of the size of an MDS, they are general and thus might become useful theoretical tools for analyzing various properties of assortative or disassortative networks.

Methods

We used network models to show the validity of the theoretical results and to perform numerical simulations. In particular, the Chung—Lu (CL) model [23] was used to generate undirected scale-free (SF) networks with N nodes and L edges (i.e., average degree $\langle k \rangle = 2L/N$). In the CL model, edges are drawn between nodes randomly selected according to node weight $(i + i_0 - 1)^\xi$, where $\xi \in [0, 1)$ and i denote the node index (i.e., $i = 1, \dots, N$), where the constant i_0 is considered to eliminate the finite-size effects (assortativity, in particular) [24]. A generated network shows that $P(k) \propto k^{-\gamma}$, where $\gamma = 1 + 1/\xi$ [23, 24] and $P(k)$ indicates the degree distribution. When generating SF networks with $\gamma < 3$, i_0 is chosen appropriately to satisfy that the maximum degree $k_{\max} < \sqrt{\langle k \rangle N}$; on the other hand, $i_0 = 1$ when $\gamma \geq 3$. When $i_0 = 1$, the CL model is equivalent to the Goh—Kahng—Kim (GKK) model [25]. In this study, we avoided the emergence of self-loops and multiple edges. Specifically, SF networks were generated using the *static.power.law.game* function in the *igraph* package (version 0.7.1) of the R software (version 3.1.1; <http://www.r-project.org>). In the case of $\gamma = \infty$, we also considered the Erdős—Rényi (ER) random networks [1], in which the node degree follows a Poisson distribution whose mean is $\langle k \rangle$.

A simulated annealing (SA) method [8] was used to adjust the degree-degree correlation, measured using the assortative coefficient r_a [4] (see also S1 File). In particular, we minimized an energy $E = |r_a - r_a^{\text{obj}}|$, where r_a^{obj} is the objective r_a , through network modifications. In this SA method, a network is modified using degree-preserving rewiring [26], in which two randomly selected edges are rewired until 1% of the edges are rewired. For example, we consider two edges, A—B and C—D, where the letters and lines are the nodes and edges, respectively. Through this edge-rewiring algorithm, the edges A—D and C—B are obtained (see Ref [26] for details). A modified network, obtained as above, is accepted with the probability $p = \min[1, \exp(E_0 - E_1)/T]$, where E_0 and E_1 are the energies obtained from the current network and the modified network, respectively, and T is the temperature. We started with $T = 1000$ and reduced T by updating as follows: $T \leftarrow 0.995 \times T$. The above procedures were continued until $E < 0.001$ or $T < 10^{-16}$.

Finding an MDS is NP-hard [21]; in particular, although it is considered as a binary integer programming problem, it is time consuming. Thus, we searched for an MDS in a network using linear programming relaxation as a binary integer programming problem, although it is possible to compute an MDS for moderate-sized networks [10, 11, 27]. We have confirmed that this relaxation does not affect the conclusions. In this study, the *lp* function in the R package *lpSolve* (version 5.6.10), a linear programming solver, was used. The size of an MDS was averaged over 100 realizations.

Results

Theoretical results

We analytically investigated the effect of degree-degree correlation (i.e., r_a) on Γ , where Γ is the MDS size. However, it was difficult to mathematically reveal the relationship between Γ and arbitrary r_a . Thus, we only evaluated the Γ of maximally assortative (respectively, maximally disassortative) networks, in which the exchange of any pair of edges does not increase (respectively, decrease) r_a (see “Assortative coefficient” in the S1 File).

The case of maximally assortative networks. Since a maximally assortative network has connections between nodes with similar degrees, it is considered as a collection of k -regular

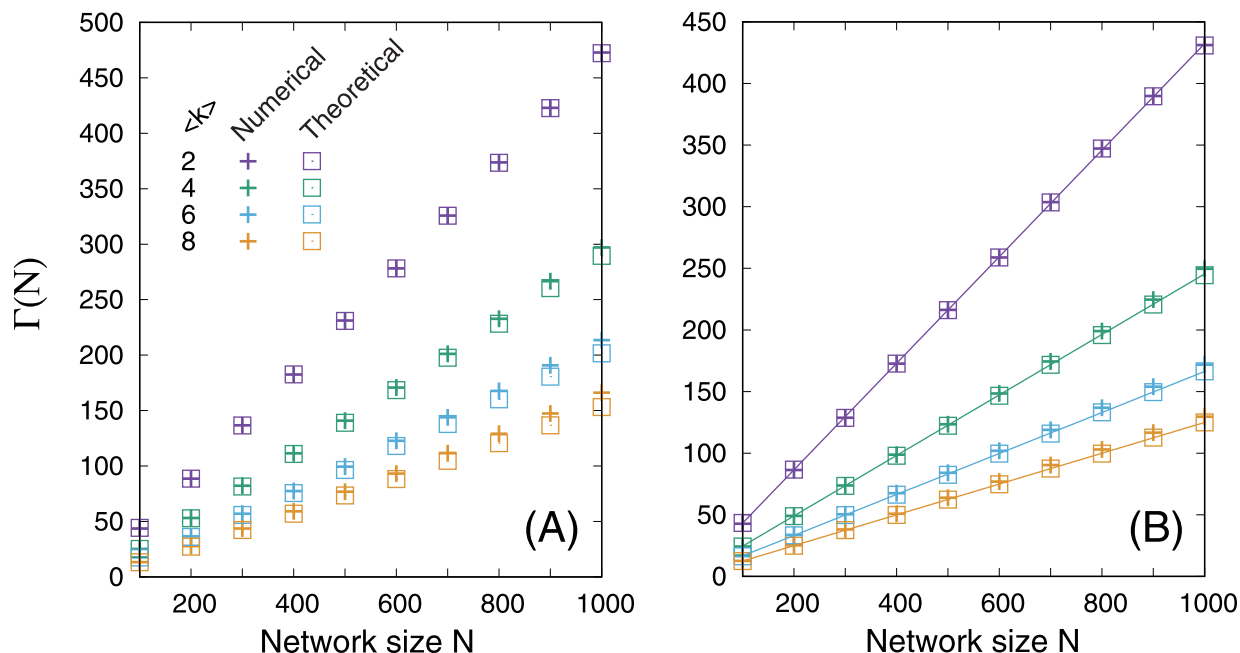


Fig 1. Network size dependency of the size of an MDS, $\Gamma(N)$, in maximally assortative networks. (A) SF networks with $\gamma = 2.5$. (B) ER random networks. The cross symbols denote the numerical results. The open squares and solid lines are the theoretical estimations; in particular, they were obtained using Eqs (1) and (2), respectively.

doi:10.1371/journal.pone.0157868.g001

random graphs ($k = k_{\min}, \dots, k_{\max}$, where k_{\min} and k_{\max} are the minimum and maximum degrees, respectively, excluding degree 0 nodes). Let n_k be the number of nodes with degree k . Since the size of the MDS, Γ , of a k -regular random graph is approximately $n_k/(k+1)$ (see “MDS in a k -regular random graph” in the [S1 File](#)), the Γ of a maximally assortative network (MAN) is derived as

$$\Gamma_{\text{MAN}} = \sum_{k=0}^{k_{\max}} \frac{n_k}{k+1}. \quad (1)$$

[Eq \(1\)](#) is in excellent agreement with the numerical results ([Fig 1](#)). Maximally assortative networks were obtained using the SA method by setting $r_a^{\text{obj}} = 1$, according to the CL and ER model networks. $\Gamma \propto N$ and the increase rate (slope) decreases with $\langle k \rangle$. This fact can be expected from the size of the MDS, $\Gamma_{\text{MAN}}^{\text{ERRG}}$, in a maximally assortative ER random graph (ERRG). $\Gamma_{\text{MAN}}^{\text{ERRG}}$ is approximately described as

$$\Gamma_{\text{MAN}}^{\text{ERRG}} = \frac{N}{\langle k \rangle} (1 - e^{-(k)}) \quad (2)$$

when $\langle k \rangle/[N-1] < 0.5$ and $N \gg 0$ (see “Size of an MDS in a maximally assortative Erdős—Rényi random graph” in the [S1 File](#)). The theoretical prediction based on [Eq \(2\)](#) is in excellent agreement with numerical results ([Fig 1B](#)); in addition to this, there is little difference in prediction accuracy between Eqs (1) and (2).

The case of maximally disassortative networks. In light of the characteristics of disassortativity (i.e., higher-degree nodes connect to lower-degree nodes), a maximally disassortative network is considered as a collection of regular random bipartite graphs (see “Maximally disassortative networks” in the [S1 File](#)). In particular, when considering an (h, k) -regular random

bipartite graph ($h > k$) $G(V_1, V_2, E)$, where each node in V_1 has degree h , each node in V_2 has degree k , and each edge connects a node in V_1 and a node in V_2 , the maximally disassortative network is represented as

$$B_{h_1, k_1} \cup B_{h_2, k_2} \cup \dots \cup B_{h_l, k_l}, \quad (3)$$

where B_{h_i, k_i} is an (h_i, k_i) -regular bipartite graph with M_{h_i} top nodes with degree h_i and N_{k_i} bottom nodes with degree k_i . Note that $h_1 \geq h_2 \geq \dots \geq h_l$ and $k_1 \leq k_2 \leq \dots \leq k_l$. Actual structures will deviate from this form to some extent.

In short, the Γ of an (h, k) -regular random bipartite graph is useful for estimating the Γ of a maximally disassortative network. Let $\alpha = m/n$, where $m = |V_1|$ and $n = |V_2|$; an upper bound of the Γ of an (h, k) -regular random bipartite graph is approximately estimated as (see “MDS in a regular random bipartite graph” in the [S1 File](#))

$$f(\beta) = \left(\frac{\alpha}{1 + k/(\alpha\beta)} + \frac{1}{1 + k\beta} \right) n, \quad (4)$$

where β is the ratio of the proportion of randomly selected nodes to the current top nodes to the proportion of randomly selected nodes to the current bottom nodes at each iteration in the recursive application of a probabilistic method (see “Size of an MDS in a maximally disassortative network” in the [S1 File](#)). The Γ of a regular random bipartite graph is obtained by minimizing [Eq \(4\)](#):

$$\Gamma_{\text{RRGB}} = g(\alpha, k, n) = \min_{\beta} f(\beta) = f(\beta_0). \quad (5)$$

The size of a maximally disassortative network is estimated by summing up $g(\alpha, k, n)$ for all decomposed regular random bipartite graphs. Since $n_k \gg n_{k'}$ is expected for $k \ll k'$, nodes (bottom nodes) with degree k may correspond to nodes (top nodes) with several degrees in each bipartite graph. We use $m(k)$ to denote the number of the top nodes corresponding to the bottom nodes with degree k , where the definition of $m(k)$ is given in the subsection “Size of an MDS in a maximally disassortative network” in the [S1 File](#). Using this $m(k)$, we redefine α by $\alpha(k) = m(k)/n_k$.

We also need to add some other factors: nodes with degree 0, and other nodes not counted by regular random bipartite graphs. To this end, we define H and K by

$$H = \max \left\{ h \mid \sum_{d=k_{\min}}^h dn_d \leq \sum_{d=h+1}^{k_{\max}} dn_d \right\} \text{ and } K = \max \left\{ k \mid \sum_{d=k_{\min}}^H dn_d \leq \sum_{d=k}^{k_{\max}} dn_d \right\},$$

respectively. Then, the size of a maximally disassortative network (MDN) is estimated as

$$\Gamma_{\text{MDN}} = n_0 + \sum_{k=k_{\min}}^H g(\alpha(k), k, n_k) + \sum_{k=H+1}^K \frac{n_k}{k+1}, \quad (6)$$

where we let $g(\alpha(k), k, n_k) = m(k)$ if $k = 1$ and $g(\alpha, k, n) = 0$ if $\alpha = 0$ or $n = 0$. The third term reflects the effect of the nodes (with degree > 0) not included in regular random bipartite graphs (“Size of an MDS in a maximally disassortative network” in the [S1 File](#)).

[Eq \(6\)](#) is in good agreement with the numerical results ([Fig 2](#)). As expected from the analysis [e.g., [Eq \(4\)](#)], $\Gamma \propto N$ and the increase rate (slope) declines with $\langle k \rangle$. However, the difference between numerical and estimated values becomes large in maximally disassortative networks of ER random networks and dense SF networks. This is because of the assumption of the theoretical estimation: we assumed that disassortative networks are completely decomposed into a collection of (h, k) -regular bipartite graphs and that the intermediate network structures of an (h, k) -regular graph are also regular, which are not necessarily true. Maximally disassortative

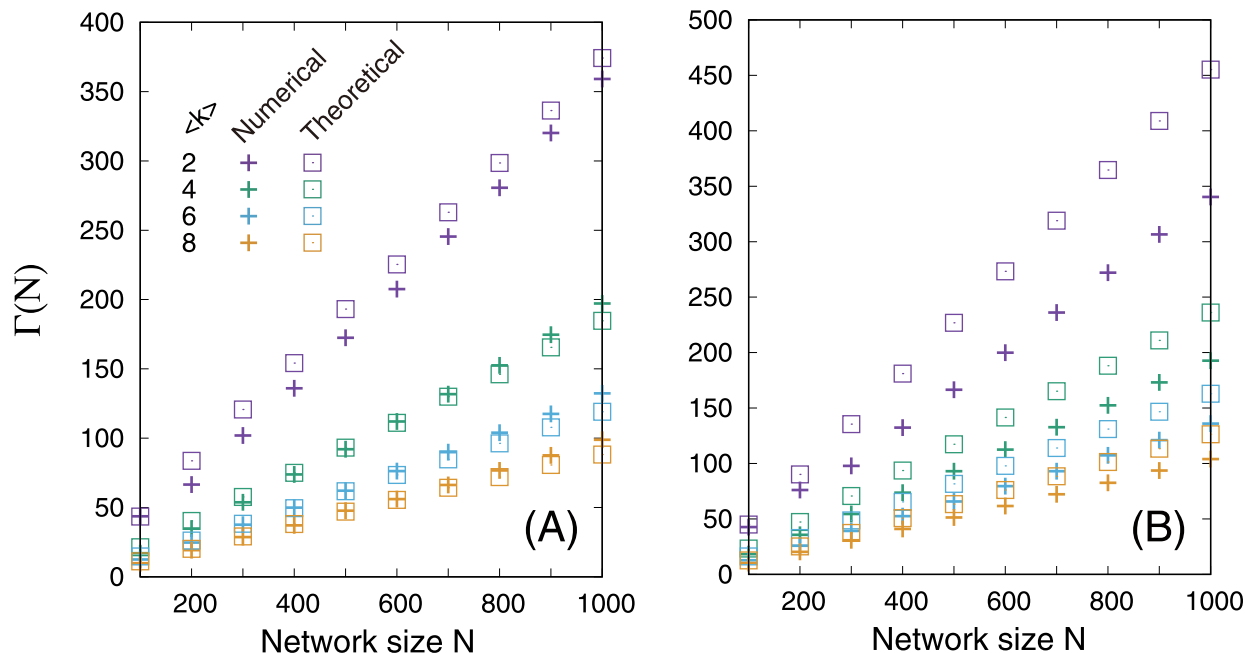


Fig 2. Network size dependency of the size of an MDS, $\Gamma(N)$, in maximally disassortative networks. (A) SF networks with $\gamma = 2.5$. (B) ER random networks. The theoretical estimations (open squares) were obtained using Eq (6).

doi:10.1371/journal.pone.0157868.g002

networks were obtained using the SA method by setting $r_a^{\text{obj}} = -1$, according to the CL model networks and ER random networks; thus, the networks may be almost maximally disassortative. This fact may cause the difference between actual and estimated values.

Numerical simulation

We performed numerical simulations to investigate the relationship between Γ and arbitrary r_a , and found that Γ declines with decreasing r_a (Fig 3), whereas it is almost independent from r_a when $r_s > 0$. Similar results were observed in both SF networks and ER random networks;

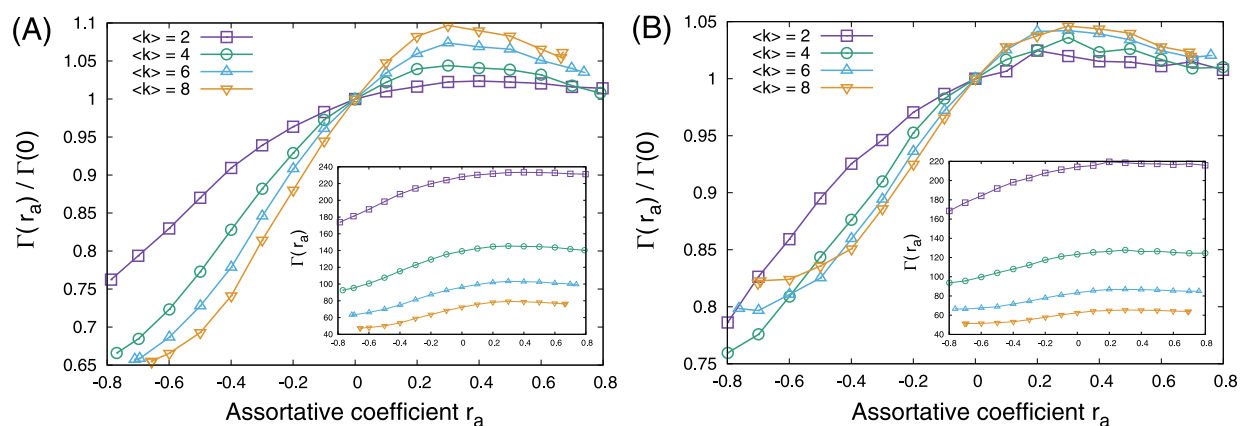


Fig 3. Relationship between the size of an MDS, $\Gamma(r_a)$, and assortative coefficient r_a in networks with $N = 500$. (A) Scale-free networks with $\gamma = 2.5$. (B) ER random networks. Note that the vertical axes of the main figure and the inset indicate relative $\Gamma(r_a)/\Gamma(0)$ and $\Gamma(r_a)$, respectively.

doi:10.1371/journal.pone.0157868.g003

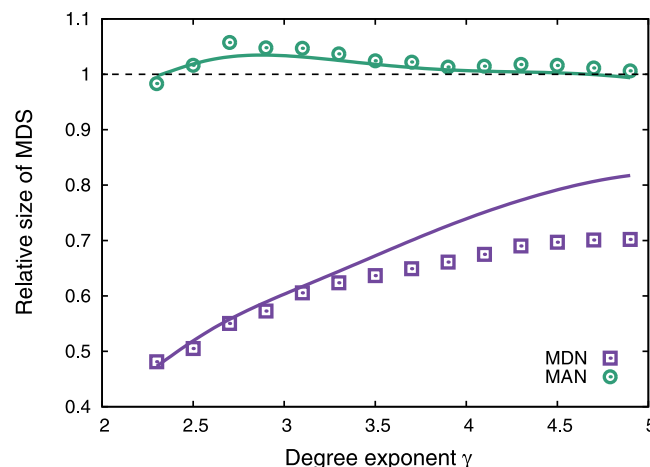


Fig 4. Degree exponent dependency of the difference between Γ_{MAN} and Γ_{MDN} in SF networks with $N = 500$ and $\langle k \rangle = 4$. The symbols denote the numerical results. The solid lines are the theoretical predictions of Γ_{MAN} (i.e., Eq (1)) and Γ_{MDN} (i.e., Eq (6)). Γ_{MAN} (Γ_{MDN}) is normalized by the size of the MDS in the uncorrelated network (i.e., network with $r_a = 0$).

doi:10.1371/journal.pone.0157868.g004

however, the effect of (dis)assortativity on Γ in SF networks is more significant than that in ER random networks.

The observed decrease in Γ with a negative r_a is due to the repulsion between high-degree nodes (i.e., hubs), which indicates disassortativity. Networks with higher-degree nodes (i.e., more heterogeneous networks) show smaller Γ [10, 11] because the trivial upper bound of Γ is described as $N - k_{\text{max}}$, where k_{max} is the highest degree in a network. In a heterogeneous network (e.g., SF network), Γ becomes smaller when hubs are separated (i.e., high-degree nodes tend to connect to low-degree nodes) because hubs independently dominate many of the other (low-degree) nodes. In assortative networks, on the other hand, since high-degree nodes tend to connect to high-degree nodes, one hub dominates the other hubs. Thus, Γ hardly decreases because many low-degree nodes need to dominate the other nodes.

These results indicate that the degree exponent γ (or degree heterogeneity) influences the difference between Γ_{MAN} and Γ_{MDN} ; in particular, they predict that more heterogeneous networks have smaller Γ_{MDN} because of the above reasons, whereas Γ_{MAN} is almost independent of γ . In fact, the difference becomes larger when γ is smaller (i.e., the degree heterogeneity becomes larger) (Fig 4). The accuracy of our theoretical estimation for Γ_{MDN} (i.e., Eq (6)) decreases with increasing γ . We discuss the reason in *Discussion and conclusions*.

Except for degree distributions, the SA method does not preserve the structural properties in networks; thus, it remains possible that the other properties mainly determine Γ . For example, clustering and modularity, which suggest community structure in complex networks [28], also influence system dynamics (e.g., epidemic dynamics [29]). To avoid this possibility, Pósfai et al. [8] evaluated these network parameters when investigating the effect of assortativity on maximum matching-based network controllability. Thus, we also investigated the relationship of Γ with clustering and modularity.

The clustering coefficient C is defined as the average ratio of the number of edges among the neighbors to the number of all possible connections among the neighbors [1]. The network modularity measure Q is defined as the fraction of edges that lie within modules rather than between modules, relative to that expected by chance [30]. We computed C and Q according to Refs. [1] and [30], respectively. In particular, the functions *transitivity* and *leading.eigenvector*.

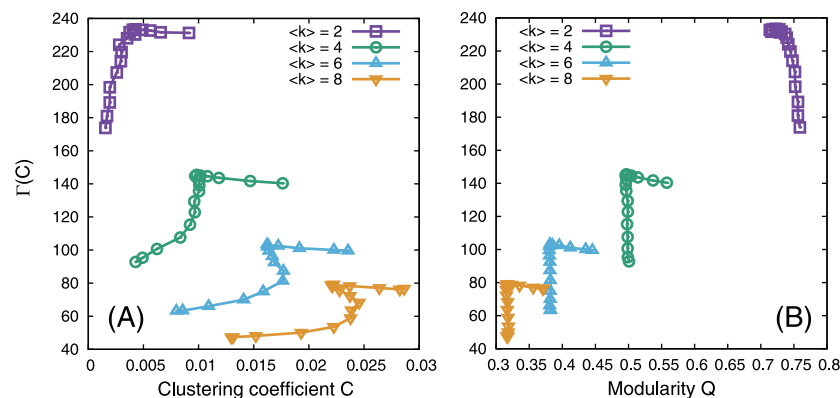


Fig 5. Correlation of the size of MDSs with the clustering coefficient (A) and modularity (B) in SF networks with $N = 500$ and $\gamma = 2.5$.

doi:10.1371/journal.pone.0157868.g005

community in the R package *igraph* (version 0.7.1) were used for calculating C and Q , respectively.

No clear relationship of Γ with C and Q was observed (Fig 5). This result indicates that r_a is the main factor in determining Γ .

Discussion and Conclusions

In summary, we showed that disassortativity reduces the size of an MDS. This finding implies that disassortativity enhances network controllability. Whalen et al. [31] showed that the presence of symmetry in a network can decrease controllability, using small nonlinear systems. Our result strongly supports their claim because assortativity can be considered as a definition of symmetry; in particular, it suggests that the impact of disassortativity on network controllability can be generally concluded. Moreover, the results suggest that biological networks are relatively easy to control, compared with social networks, because biological networks (e.g., protein interaction networks) are disassortative, whereas social networks are assortative [4]. Liu et al. [3] suggested the difficulty in controlling social networks because it is necessary to control most individuals separately to control the whole system. Their observation and suggestion may be consistent with our theoretical predictions.

This result does not contradict a previous study [8], which showed that (dis-)assortativity does not increase network controllability in terms of structural controllability [3] because the model of control adopted in the MDS-based approach is different from the standard one. A relationship between an MDS and structural controllability is discussed in Ref [10], which states that, if every edge in a network is bi-directional and every node in an MDS can control all of its outgoing edge values independently, then the network is structurally controllable under a linear dynamic model. It should be noted that this MDS-based model employs a much stronger assumption than that used by the standard structural controllability model [3] in which only the values of driver nodes can be controlled. Thus, the MDS-based model may not be directly applicable to the control of typical dynamic systems. However, as mentioned in the *Introduction*, the MDS is a well-known concept in graph theory and has been applied to the design and control of various engineering systems. In addition, it is reported in several studies that nodes in an MDS and its variants tend to have important biological roles [9, 14–16]. For example, Wuchty [14] showed that proteins in an MDS are enriched with cancer-related and virus-targeted genes in protein interaction networks; furthermore, they found that such proteins have a higher impact on network resilience.

However, more careful examinations may be required to conclude the effect of degree-degree correlation on the size of an MDS. Specifically, we may need to consider the effect of other network properties on the size of an MDS. For example, Bianchin et al. [32] showed the importance of diameter in the controllability of complex networks, although the definition of network controllability in Ref [32] is different from that in our study. Disassortative networks may show a larger diameter than random networks do because of the repulsion between high-degree nodes.

Recent studies pointed out that the minimum number of driver nodes is not necessarily appropriate as a measure of the degree of controllability because the use of a few driver nodes may need a forbiddingly high amount of energy [32–36]. Instead, these studies focus on the energy needed to control network systems. In general, the minimum and maximum energies can be determined by evaluating the eigenvalues of the Gramian matrix: the minimum and maximum normalized energy costs are given by the reciprocal of the maximum and minimum eigenvalues [33]. By analyzing the Gramian matrix, researchers were able to perform various studies on the control energy, which include scaling laws with respect to control time [33], trade-off between the control energy and the number of control nodes [34], relationships between the control energy and the network diameter [32], algorithms for selecting a set of minimal control nodes under a given energy constraint [35], and relationships between the control energy and the number of control nodes in scale-free networks [36].

In particular, Yan et al. showed that, for a linear and continuous network control model with the power-law degree distribution $P(k) \propto k^{-\gamma}$ and self-loops, the maximum control energy scales sublinearly as $N^{1/(\gamma-1)}$ when $N_D = N$, whereas it scales exponentially as e^N when $N_D = 1$, where N and N_D denote the number of network nodes and driver nodes, respectively [36]. In the MDS-based control model, it is assumed that each node in an MDS can control itself and its connecting edges independently [10]. On the other hand, by the definition of the MDS, each node in the network must be in an MDS or have at least one edge connecting to a node in the MDS. Therefore, it is meant that each node has an independent control input (if multiple control inputs are assigned to a node, we can use only the one that gives 0 values to the other inputs). This corresponds to the case of $N_D = N$ in Ref [36], which suggests that the MDS-based control model requires relatively low energy costs, although it must control edges connecting to nodes in an MDS independently.

We proposed a method for estimating the size of an MDS in maximally (dis-)assortative networks from degree distributions. Although the estimation of MDS size or domination number is a central topic in graph theory, it has been limited in the context of random graphs without degree-degree correlations such as ER random networks [37], k -regular random graphs [38], and Barabási-Albert networks [39]. Our analytical approach is applicable to the estimation of the domination number of a network with a given degree distribution. However, our theoretical estimations (i.e., Eqs (1) and (6)) have limitations. In particular, the estimation in ER random networks is better than that in SF networks in the maximally assortative case; on the other hand, the opposite tendency was observed in the maximally disassortative case. This is because of the difference in degree heterogeneity between ER random networks and SF networks. The theoretical estimation in the maximally assortative case (i.e., Eq (1)) assumes a collection of k -regular random graphs. The assumption of k -regular random graphs is appropriate in ER random networks because node degree is almost similar. For a large k , however, this assumption is hardly satisfied in SF networks because of the existence of a few high-degree nodes (i.e., hubs) due to heterogeneous (power-law) degree distributions. Thus, Eq (1) shows a lower prediction accuracy in SF networks. On the other hand, the theoretical estimation in the maximally disassortative case (i.e., Eq (6)) considers the assumption of a group of (h, k) -regular random bipartite graphs. This assumption is not appropriate in ER random

networks when the difference between h and k is large (i.e., the maximally disassortative case), although it is relatively suitable in SF networks because of the hubs. In particular, it is difficult to obtain (h, k) -regular random bipartite graphs in such a case (i.e., $h \gg k$) because node degree is almost similar in ER random networks; thus, [Eq \(6\)](#) exhibits a lower accuracy.

In this study, we focused on the MDS in maximally assortative and disassortative networks. However, the proposed methodology may be extended to the general degree-degree correlation case by making use of a weighted mixture of (h, k) -regular bipartite networks. The proposed methodology (i.e., combination of network decomposition and recursive application of a probabilistic method) may also be applied to the analysis of other graph theoretic properties (e.g., maximum matching, independent set) in assortative/disassortative networks. Our study could lead to a deeper understanding of network controllability and other properties of complex networks.

Supporting Information

S1 File. Supplementary information.
(PDF)

Author Contributions

Conceived and designed the experiments: KT TA. Performed the experiments: KT TA. Analyzed the data: KT. Wrote the paper: KT TA.

References

1. Albert R, Barabási A-L. Statistical mechanics of complex networks. *Rev Mod Phys.* 2002; 74(1): 47–97. doi: [10.1103/RevModPhys.74.47](https://doi.org/10.1103/RevModPhys.74.47)
2. Takemoto K, Oosawa C. Modeling for evolving biological networks. in: Dehmer M, Basak SC, editors. *Statistical and Machine Learning Approaches for Network Analysis*. New York: John Wiley & Sons; 2012. pp. 77–108.
3. Liu Y, Slotine J, Barabási A-L. Controllability of complex networks. *Nature* 2011; 473(7346): 167–173.
4. Newman MEJ, Assortative mixing in networks. *Phys Rev Lett.* 2002; 89(20): 208701. doi: [10.1103/PhysRevLett.89.208701](https://doi.org/10.1103/PhysRevLett.89.208701) PMID: [12443515](https://pubmed.ncbi.nlm.nih.gov/12443515/)
5. Egúíluz VM, Klemm K. Epidemic threshold in structured scale-free networks. *Phys Rev Lett.* 2002; 89(10) 108701. doi: [10.1103/PhysRevLett.89.108701](https://doi.org/10.1103/PhysRevLett.89.108701) PMID: [12225235](https://pubmed.ncbi.nlm.nih.gov/12225235/)
6. Wang B, Zhou T, Xiu ZL, Kim BJ. Optimal synchronizability of networks. *Eur Phys J B.* 2007; 60(1):89–95. doi: [10.1140/epjb/e2007-00324-y](https://doi.org/10.1140/epjb/e2007-00324-y)
7. Hasegawa T, Takaguchi T, Masuda N. Observability transitions in correlated networks. *Phys Rev E.* 2013; 88(4): 042809. doi: [10.1103/PhysRevE.88.042809](https://doi.org/10.1103/PhysRevE.88.042809)
8. Pósfai M, Liu Y-Y, Slotine J-J, Barabási A-L. Effect of correlations on network controllability. *Sci Rep.* 2013; 3: 1067.
9. Milenković T, Memišević V, Bonato A, Pržulj N. Dominating biological networks. *PLoS One.* 2011; 6(8): e23016. doi: [10.1371/journal.pone.0023016](https://doi.org/10.1371/journal.pone.0023016) PMID: [21887225](https://pubmed.ncbi.nlm.nih.gov/21887225/)
10. Nacher JC, Akutsu T. Dominating scale-free networks with variable scaling exponent: Heterogeneous networks are not difficult to control. *New J Phys.* 2012; 14(7): 073003. doi: [10.1088/1367-2630/14/7/073005](https://doi.org/10.1088/1367-2630/14/7/073005)
11. Nacher JC, Akutsu T. Analysis on controlling complex networks based on dominating sets. *J Phys.* 2013; 410: 012104.
12. Molnár Jr F, Sreenivasan S, Szymanski BK, Korniss G. Minimum dominating sets in scale-free network ensembles. *Sci Rep.* 2013; 3: 1736.
13. Molnár Jr F, Derzsy N, Czabarka É, Székely L, Szymanski BK, Korniss G. Dominating scale-free networks using generalized probabilistic methods. *Sci Rep.* 2014; 4: 6308.
14. Wuchty S. Controllability in protein interaction networks. *Proc. Natl. Acad. Sci. USA.* 2014; 111(19): 7156–7160. doi: [10.1073/pnas.1311231111](https://doi.org/10.1073/pnas.1311231111) PMID: [24778220](https://pubmed.ncbi.nlm.nih.gov/24778220/)

15. Khuri S, Wuchty S. Essentiality and centrality in protein interaction networks revisited. *BMC Bioinformatics*. 2015; 16(1): 109. doi: [10.1186/s12859-015-0536-x](https://doi.org/10.1186/s12859-015-0536-x) PMID: [25880655](https://pubmed.ncbi.nlm.nih.gov/25880655/)
16. Zhang X-F, Ou-Yang L, Zhu T, Wu M-Y, Dai D-Q. Determining minimum set of driver nodes in protein-protein interaction networks. *BMC Bioinformatics*. 2015; 16(1): 146. doi: [10.1186/s12859-015-0591-3](https://doi.org/10.1186/s12859-015-0591-3) PMID: [25947063](https://pubmed.ncbi.nlm.nih.gov/25947063/)
17. Nepusz T, Vicsek T. Controlling edge dynamics in complex networks. *Nat Phys*. 2012; 8(7): 568–573.
18. Alzoubi K, Wan PJ, Frieder O. Message-optimal connected dominating sets in mobile ad hoc networks. *Proc. 3rd Int. Symp. on Mobile ad hoc Networks and Computing*. 2002: 157–164.
19. Stojmenovic I, Seddigh M, Zunic J. Dominating sets and neighbor elimination-based broadcasting algorithms in wireless networks. *IEEE Trans. Parallel Distrib. Syst.* 2012; 13(1): 14–25. doi: [10.1109/71.980024](https://doi.org/10.1109/71.980024)
20. Cokuslu D, Erciyes K, Dagdeviren O. A dominating set based clustering algorithm for mobile ad hoc networks. *Proc. Int. Conf. on Computational Science (Lecture Notes in Computer Science vol. 3991)*. 2006: 571–578.
21. Haynes TW, Hedetniemi ST, Slater PJ. *Fundamentals of domination in graphs*. New York: Chapman and Hall/CRC; 1998.
22. Sampath A, Thampi SM. An ACO algorithm for effective cluster head selection. *J Adv Inf Technol*. 2011; 2(1): 50–56.
23. Chung F, Lu L. Connected components in random graphs with given expected degree sequences. *Ann Comb*. 2002; 6(2): 125–145. doi: [10.1007/PL00012580](https://doi.org/10.1007/PL00012580)
24. Cho YS, Kim JS, Park J, Kahng B, Kim D. Percolation transitions in scale-free networks under the Achlioptas process. *Phys Rev Lett*. 2009; 103(13): 135702. doi: [10.1103/PhysRevLett.103.135702](https://doi.org/10.1103/PhysRevLett.103.135702) PMID: [19905523](https://pubmed.ncbi.nlm.nih.gov/19905523/)
25. Goh K-I, Kahng B, Kim D. Universal behavior of load distribution in scale-free networks. *Phys Rev Lett*. 2001; 87(27): 278701. doi: [10.1103/PhysRevLett.87.278701](https://doi.org/10.1103/PhysRevLett.87.278701) PMID: [11800921](https://pubmed.ncbi.nlm.nih.gov/11800921/)
26. Maslov S, Sneppen K. Specificity and stability in topology of protein networks. *Science*. 2002; 296(2002): 910–913. doi: [10.1126/science.1065103](https://doi.org/10.1126/science.1065103) PMID: [11988575](https://pubmed.ncbi.nlm.nih.gov/11988575/)
27. Zhao J-H, Habibulla Y, Zhou H-J. Statistical mechanics of the minimum dominating set problem. *J Stat Phys*. 2015; 159: 1154–1174. doi: [10.1007/s10955-015-1220-2](https://doi.org/10.1007/s10955-015-1220-2)
28. Fortunato S. Community detection in graphs. *Phys Rep*. 2010; 486(3–5): 75–174. doi: [10.1016/j.physrep.2009.11.002](https://doi.org/10.1016/j.physrep.2009.11.002)
29. Ritchie M, Berthouze L, House T, Kiss IZ. Higher-order structure and epidemic dynamics in clustered networks. *J Theor Biol*. 2014; 348: 21–32. doi: [10.1016/j.jtbi.2014.01.025](https://doi.org/10.1016/j.jtbi.2014.01.025) PMID: [24486653](https://pubmed.ncbi.nlm.nih.gov/24486653/)
30. Newman MEJ. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA*. 2006; 103(23): 8577–8582. doi: [10.1073/pnas.0601602103](https://doi.org/10.1073/pnas.0601602103) PMID: [16723398](https://pubmed.ncbi.nlm.nih.gov/16723398/)
31. Whalen AJ, Brennan SN, Sauer TD, Schiff SJ. Observability and controllability of nonlinear networks: The role of symmetry. *Phys Rev X*. 2015; 5: 011005.
32. Bianchin G, Pasqualetti F, Zampieri S. The role of diameter in the controllability of complex networks. *IEEE Conf. on Decision and Control*. 2015: 980–985.
33. Yan G, Ren J, Lai Y-C, Lai C-H, Li B. Controlling complex networks: How much energy is needed? *Phys Rev Lett*. 2012; 108(21): 218703. PMID: [23003312](https://pubmed.ncbi.nlm.nih.gov/23003312/)
34. Pasqualetti F, Zampieri S, Bullo F. Controllability metrics, limitations and algorithms for complex networks. *IEEE Trans Control Netw Syst*. 2014; 1(1): 40–52.
35. Tzoumas V, Rahimian MA, Pappas GJ, Jadbabaie A. Minimal actuator placement with bounds on control effort. *IEEE Trans Control Netw Syst*. *in press*.
36. Yan G, Tsekenis G, Barze B, Slotine J-J, Liu Y-Y, Barabási A-L. Spectrum of controlling and observing complex networks. *Nat Phys*. 2015; 11: 779–786.
37. Henning MA, Yeo A. The domination number of a random graph. *Util Math*. 2014; 94: 315–328.
38. Molloy M, Reed B. The dominating number of a random cubic graph. *Random Struct Algorithms*. 1995; 7(3): 209–222. doi: [10.1002/rsa.3240070303](https://doi.org/10.1002/rsa.3240070303)
39. Cooper C, Klasing R, Zito M. Lower bounds and algorithms for dominating sets in web graphs. *Internet Math*. 2005; 2(3): 275–300. doi: [10.1080/15427951.2005.10129105](https://doi.org/10.1080/15427951.2005.10129105)